THIRD QUARTERLY REPORT

FOR

THE DEVELOPMENT OF COMPUTATIONAL TECHNIQUES

FOR THE IDENTIFICATION OF

LINEAR AND NONLINEAR MECHANICAL SYSTEMS SUBJECT TO RANDOM EXCITATION

Contract No. NAS5-10106

Prepared by

Will Gersch

Midwest Applied Science Corp.
McClure Research Park
1205 Kent Avenue
West Lafayette, Indiana

September 1966

For

Goddard Space Flight Center
Greenbelt, Maryland

TABLE OF CONTENTS

ON THE IDENTIFICATION

OF

LINEAR AND NONLINEAR STRUCTURAL SYSTEMS

Objective: The development of computational techniques for the identification of linear and nonlinear mechanical systems subject to random excitation.

Summary: Computational procedures have been suggested to determine the differential equation governing the motion of linear and nonlinear structural systems subject to random excitation when the system excitation and response are observed. The objective of this effort is to yield the transfer functions, impedances and damping coefficients of linear systems as well as to determine the nonlinearities in the spring and damping coefficients governing the motion of nonlinear structures.

In general the computational procedure employed for the identification of the unknown structure consists of three stages. The first is the generation of model reference hypotheses concerning the number of degrees of freedom of the system. The second stage is one of parameter estimation in which the assumed model is fit to the observed data. The final stage consists of a verification of the validity of the assumed model. It

therefore involves the statistical inference procedures of hypotheses testing.

In the first quarterly progress report (1) a quasilinearization--least squares--recursive smoothing procedure to accomplish the parameter estimation stage of the identification procedure computations was described. Theoretically this procedure is sufficiently general to accomplish the parameter estimation for both linear and nonlinear systems and preliminary computation examples were illustrated. Difficulty in getting quasilinear computational solutions to converge to a correct solution when the initial guess was excessively far from the correct solution was experienced. This difficulty motivated examination of an alternate least squares identification procedure that is simpler to implement but is only suitable for the identification of linear systems subject to a zero mean random excitation. The theory underlying the least squares identification scheme for randomly excited linear systems was described in the last quarterly progress report (2).

During the current quarterly interval, effort has continued to be concentrated on the linear system identification procedures. In the preceding report, (2), the identification of the parameters of an unknown linear dynamical system was reduced to the identification of the unknown parameters in a stochastic difference equation or autoregressive scheme [Equation (25), Section 3.2.2]. In this report, the theory leading to the autoregressive scheme representation of the unknown parameters of the linear dynamical system is briefly reviewed as are the statistical results on the estimation of the parameters and the determination of the order of autoregressive schemes. (The order, k, of the autoregressive model introduced is two times the number of degrees of freedom of the original, unknown continuous dynamic system.) A computational example illustrates some of the material discussed on the autoregressive scheme. It is anticipated that more extensive computational experiments on linear system identification will be conducted in the next quarterly interval and that the investigation of procedures for the identification of nonlinear systems will also be resumed.

1. INTRODUCTION

The objective of the investigation is to develop a computational procedure for the identification of mechanical structures that are driven by a random excitation. In particular, the structures can be conceived of as an arbitrary collection of lumped spring-mass-damper coefficients in the linear case, or by a polynomial description of the nonlinearities in the nonlinear case.

The approach employed for the identification of the unknown structure consists of 3 stages. The first is the generation of hypotheses concerning the number of degrees of freedom of the system and the form of the nonlinearities. In effect, this prescribes a conceptual and computational model for the system. In the second stage, the observed data, corresponding to the excitation and response of the system, is used to determine parameters or coefficients of the model assumed to represent the system. The final stage consists of a verification of the validity of the assumed computational model. This is to be accomplished by comparing the response of the system model to the response of the actual system. Subject to an "energy" response criterion, the assumed model is either accepted or an alternative model is assumed and computed on. In case of the latter alternative; the procedure is iterated, starting once again with stage 1.

In the first quarterly progress report (1), a quasi-linearization-least squares sequential estimation procedure, suitable for the identification of both linear and nonlinear systems was discussed and some computational examples were given. The approach may be thought of as the identification of an unknown system by comparison with a sequence of model reference conjectures. The quasilinearization procedure is suitable for deterministic (swept sine wave for example) and random excitation driving forces.

In the second quarterly progress report (2), a least squares procedure for the identification of linear time invariant systems under zero mean random force excitation and regularly spaced observations was introduced. As a consequence, the problem of identifying the parameters of an unknown k/2 degree of freedom dynamical system was found to be equivalent to the problem of estimating the unknown parameters in a k^{th} order autoregressive scheme.

The asymptotic statistical properties of the autoregressive parameter estimation procedure have been demonstrated to be equivalent to the results in ordinary regression theory (Mann and Wald 1943, Reference (3)) and have been extensively studied and reported on since (4,5).

In this report we briefly review the mathematical basis for our linear system identification procedure and also the review of mathematical results associated with autoregressive models (the estimation of the parameters and the determination of the order of the autoregressive scheme).

The heuristic energy fit criterion suggested earlier to determine the suitability of the model assumed to fit to the unknown linear system is shown to be equivalent to the residual variance/observed power, statistic used to determine the order of an autoregressive scheme.

A preliminary example of a least squares fit to an autoregressive scheme is illustrated. It is anticipated that more extensive computational experiments on linear system identification will be conducted during the next quarterly interval.

2. ANALYSIS

2.1 Objective

Our objective is to describe a computational procedure which will permit identification of a continuous parametrically described unknown stationary linear dynamic system excited by white noise which is observed by a regular sampling process. The situation is depicted in Figure 1.

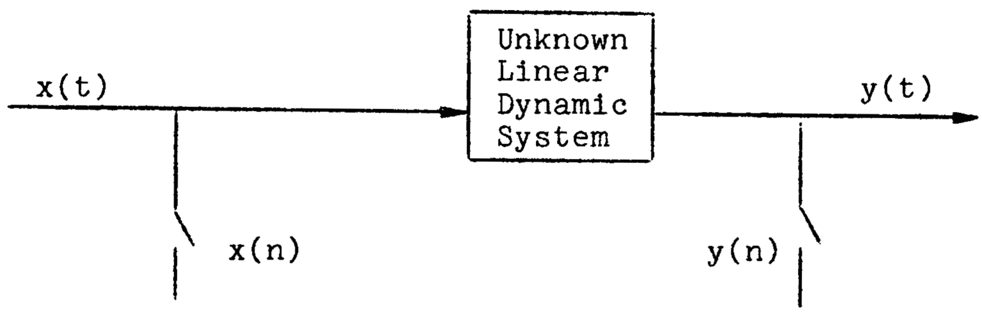

FIGURE 1. The Identification Problem Considered

The unknown system is assumed to be represented by the dynamic equations

$$\dot{q}(t) = A\,q(t) + b\,x(t) \tag{1}$$

$$y(t) = c'\,q(t)$$

$$A = \left[\begin{array}{c|ccc} 0 & & & \\ \vdots & & I & \\ 0 & & & \\ \hline -a_k & -a_{k-1} & & -a_1 \end{array}\right] \; ; \; b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \; ; \; c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

In the matrix A, the parameters $k, a_1, \ldots a_k$ are unknown and in addition the time function $x(t)$ is assumed to be a sample function of a white noise process.

The system input $x(t)$ and its response are regularly sampled over a finite time observation interval (for the purpose of digital computation) and consequently give rise to the observed time sequences $x(n)$, $y(n)$; $n=1,2,\ldots m$.

Our requirement is that we estimate the unknown system parameters $k, a_1, \ldots a_k$ from the finite duration time series $x(n)$ a $y(n)$; $n=1,2,\ldots m$. From this knowledge we may compute the linear system transfer function, impedance, etc.

2.2 Outline of the Procedure

The regularly sampled system (2.1.1) can be expressed as the discrete time system

$$q(n+1) = F\, q(n) + f\, \omega(n); \qquad (1)$$

$$y(n) = c'\, q(n)$$

where the kxk matrix F and the kx1 column vector f are functions of the unknown system parameters $k, a_1, \ldots a_k$ and $\omega(n)$ is a white noise sequence. In the preceding progress report it was demonstrated that in general

$$F = \exp\,(TA) \qquad (2)$$

where T is the sampling interval and f is in general a more complicated function of the parameters $a_1, \ldots a_k$. It should be noted that the representation in (2) is an equivalent of (2.1.1) from the point of view of the identification problem in the sense that it exposes the system parameters sufficiently to permit them to be estimated*

*This equivalence is distinctively different from that achieved in the more usual discrete representation of continuous linear systems. The latter problem is well discussed by Blackman (6), the former problem is not known to have been treated in the literature. The distinction between our representation and the more usual one is a consequence of the fact that the usual time discrete representation of a continuous time system or process is an approximation and the discrete time series may be of arbitrary order depending upon the quality of the approximation. The choice of our representation (1) is motivated by several results in the identification of sampled data systems, (R.C.K. Lee (7)) and the fact that the white noise signal source can be employed.

The equivalence of (1) and (2.1.1) is reviewed in Section (2.3).

In Section (2.4) it is demonstrated that the system (1) can be put into the form

$$y(n) = \sum_{i=1}^{k} -\alpha_i \, y(n-i) + \sum_{i=1}^{k} \beta_i \, \omega(n-i); \tag{3}$$

$$n = 1,2,\ldots m$$

Equation (3) is in the form of a mixed autoregressive-moving average model, (Hannan [4]). The parameters $\{\alpha_i\}$ and $\{\beta_i\}$ are functions of the unknown system parameters $\{a_i\}$. We replace the second series in (3) by an equivalent autocorrelated series, $\eta(n)$, where

$$\mu(n) = \sum_{i=1}^{k} \beta_i \, \omega(n-i) \tag{4}$$

and are left with the k^{th} order autoregressive model

$$y(n) = -\alpha_i \, y(n-i) - \cdots - \alpha_k \, y(n-k) + \eta(n); \tag{5}$$

$$n = 1, \cdots, m.$$

The unknown coefficients $\alpha_1, \ldots \alpha_k$ in the autoregressive model are estimated by a least squares procedure and are subsequently transformed to the unknown system parameters $a_1, \ldots a_k$.

The following is a list of the analytic steps employed in the procedure and the corresponding sections in which they are discussed.

(i) The equivalent continuous time and discrete time models (2.3).

(ii) The transformation between the discrete time model and the autoregressive model (2.4).

(iii) The transformation between the $\{\alpha_i\}$ and the $\{a_i\}$ (5).

(iv) The estimation of $k, \alpha_1, \ldots \alpha_k$ from the autoregressive model (2.6).

Items (ii) and (iv) follow respectively from adaptions from the work of R. C. K. Lee (7) and E. J. Hannan (4). Items (i) and (iii) are not known to have explicitly appeared before.

In addition, the digital computer programs written to accomplish the estimation of the coefficients and some numerical results in the autoregressive model are discussed in Section (2.7).

2.3 The Equivalent Time Continuous and Time Discrete Models

The regularly sampled versions of the continuous time system signals $x(t)$ and $y(t)$ (2.1.1) give rise to the discrete time series $x(n)$ and $y(n)$ from which we wish to estimate the unknown continuous system parameters. The discrete versions of the input-output relationships of a linear time invariant system can be thought of either as a discrete time-time invariant system of equivalently as a mixed model moving average-autoregressive model. If our parametricized representation of the unknown continuous time system were represented as an autoregressive model we could employ the techniques of regression analysis to estimate the unknown coefficients in the autoregressive model and subsequently transform these back to the continuous system parameters. Hence we are motivated to seek an autoregressive model equivalent of the continuous time system. For simplicity, let's arbitrarily consider one which is the same order as the number of state variables in the original unknown dynamic system. (The variable k, corresponding to $k/2$ d.odf. system.) Since the estimation of parameters in the autoregressive model is a consequence of the structure of the covariance properties of the "system response", we examine the covariance properties of the k state time continuous and k state time discrete models.

Consider the state variable representation of a linear dynamical system in the form (2.1.1)

$$q(t) = A\,q(t) + b\,x(t);$$

$$y(t) = c'\,q(t). \qquad (1)$$

Our interest is in the covariance stationary properties of $y(t)$ hence we consider the stationary or steady state solution

$$q(t) = \int_{-\infty}^{t} e^{(t-\lambda)A}\,bx(\lambda)d\lambda$$

$$= \int_{0}^{\infty} e^{\lambda A}\,bx(t-\lambda)\,d\lambda. \qquad (2)$$

Since $E\,x(t) = 0$; $E\,q(t) = 0$ and consequently the covariance matrix associated with the state variable $q(t)$ is

$$E\,q(t)\,q'(t-\tau) = E\int_{0}^{\infty}\int_{0}^{\infty} e^{\lambda A}\,bx(t-\lambda)\,x(t-\tau-\mu)b'e^{\mu A'}d\lambda d\mu$$

$$= \int_{0}^{\infty}\int_{0}^{\infty} e^{\lambda A}b\,\delta(\tau+\mu-\lambda)b'\,e^{\mu A'}\,d\lambda d\mu$$

$$= \int_{0}^{\infty} e^{(\tau+\mu)A}\,bb'\,e^{\mu A'}\,d\mu \qquad (3)$$

$$= e^{\tau A}M$$

where

$$M = \int_0^\infty e^{\mu A}\, bb' e^{\mu A'} d\mu \tag{4}$$

and M is positive definite. Correspondingly, the covariance of $y(t)$ is

$$E\, y(t)y(t-\tau) = c' e^{\tau A} M c \tag{5}$$

which can be evaluated for any matrix A.

Now consider the k^{th} order discrete time dynamic system

$$\begin{aligned} q(n+1) &= F\, q(n) + f\, x(n); \\ y(n) &= d'\, q(n) \end{aligned} \tag{6}$$

where F, f and d are respectively $k \times k$, $k \times 1$ and $k \times 1$ matrices and $x(n)$, $n = 0,1,2\ldots$ is a zero mean independent gaussian distributed sequence with variance σ^2 (a discrete version of $x(t)$).

To determine the covariance $E\, y(n)y(m)$, first consider the solution $q(n)$. From (6)

$$\begin{aligned} q(n+2) &= F\, q(n+1) + f\, x(n+1) \\ &= F^2\, q(n) + Ff\, x(n) + f\, x(n+1) \end{aligned} \tag{7}$$

$$\begin{aligned} q(n+m) &= F^m\, q(n) + F^{m-1} f\, x(n) + F^{m-2} f\, x(n+1) + \cdots f x(n+m-1) \\ &= F^m\, q(n) + \sum_{k=o}^{m-1} F^{m-1-k} f\, x(n+k) \end{aligned}$$

Considering only the steady state part of the solution we have that

$$E\, q(n+1)\, q'(n\underline{+}m) = E\, fx(n) \sum_{k=o}^{m-1} x(n+k)f'(F^{m-1-k})'$$
$$= f \sum_{k=o}^{m-1} Ex(n)x(n+k)f'(F^{m-1-k})' \tag{8}$$

Since $E\, x(n) = 0$; $E\, x(n)x(n+k) = 0$ for $k \neq 0$.

$$E\, q(n+1)q'(n+m) = ff'\, (F^{m-1})' = (F^{m-1})\, ff' \tag{9}$$

In (9) we used the symmetry property of the covariance matrix and the fact that the matrix ff' is also symmetric. Therefore

$$E\, y(n)y(n+m) = d'\, (F^m)ff'\, d \tag{10}$$

If we identify

$$F = \exp TA; \quad mT = \tau$$
$$d = c \tag{11}$$
$$ff' = M$$

the covariance of the time discrete model (10) is identical at the lag points $mT = \tau$, $m = 0,1,...$ to the covariance of the time continuous model, and this holds for arbitrary T.

Now M can be explicitly determined for any A and from M we could determine ff'.

We know that the time discrete state variable model (6) can be put into the autoregressive form (see Section 2.4)

$$y(n) = \sum_{i=1}^{k} -\alpha_i \; y(n-i) + \sum_{i=1}^{k} \beta_i \; x(n-i) \tag{12}$$

where the α_i are only functions of the parameters in F. Also we can write (12) in the form

$$y(n) = \sum_{i=1}^{k} -\alpha_i \; y(n-i) + \eta\,(n) \tag{13}$$

where $\mu(n)$ is an autocorrelated series derived from the moving average component in (13).

$$\mu(n) = \sum_{i=1}^{k} \beta_i \; x(n-i) \tag{14}$$

We can employ a least squares parameter estimate to determine the α_i (see Section 2.6). Consequently we conclude that from the point of view of the estimation of the unknown coefficients in the linear dynamic system in (1) we can employ the model

$$\begin{aligned} q(n+1) &= F\, q(n) + f\, x(n) \\ y(n) &= c\, q(n) \end{aligned} \tag{15}$$

where $F = \exp TA$, and f is an unknown $k \times 1$ column vector.

2.4 The Transformation Between the Discrete Time and the Autoregressive Model

In this section we demonstrate that the discrete time dynamic system

$$q(n+1) = F\,q(n) + f\,x(n)$$
$$y(n) = c'\,q(n) \tag{1}$$

where F,f and d are respectively kxk, kx1 and kx1 matrices and $x(n)$, $n = 0,1,\ldots$ is a zero mean independent gaussian distributed sequence, can be written in the form of a mixed autoregressive-moving average model

$$y(n) = \sum_{i=1}^{k} -\alpha_i\, y(n-i) + \sum_{i=1}^{k} \beta_i\, x(n-i) \tag{2}$$

Under the nonsingular transformation

$$s = B\,q; \tag{3}$$

where

$$B = \begin{bmatrix} c' \\ c'F \\ \vdots \\ c'F^{n-1} \end{bmatrix} \tag{4}$$

it can be shown by direct substitution, that (1) is transformed into the canonical form (see Quarterly Progress Report #2, Appendix, for details of this demonstration).

$$s(n+1) = \Phi\, s(n) + d\, x(n) \tag{5}$$

$$y(n) = c'\, s(n)$$

where

$$\Phi = \left[\begin{array}{c|ccc} 0 & & I & \\ \vdots & & & \\ 0 & & & \\ \hline -\alpha_k & -\alpha_{k-1} & & -\alpha_1 \end{array}\right] ; \; c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} ; \; s(n) = \begin{bmatrix} s_1(n) \\ \vdots \\ s_k(n) \\ \end{bmatrix} \tag{6}$$

and d is some $n x l$ column vector.

That the canonical form (5) can be represented in the form (2) can also be demonstrated by direct substitution. From (5)

$$s_j(n+1) = s_{j+1}(n) + d_j\, x(n) \quad j = 1,2,\ldots k-1 \tag{7}$$

and

$$s_k(n+1) = \sum_{i=1}^{k} -\alpha_{k+1-i}\, s_i(n) + b_k\, x(n). \tag{8}$$

We observe that from (5)

$$y(n) = q_1(n) \tag{9}$$

Therefore we solve (7) for $s_j(n)$ in terms of $s_1(n+j-1)$ and obtain

$$s_j(n) = s_1(n+j-1) - \sum_{i=1}^{j-1} d_i\, x(k-(j-1)-i); \; j=2,\cdots k \tag{10}$$

Substituting (10) and (9) into (8) yields

$$y(n+k) = \sum_{i=1}^{k} -\alpha_i \; y(n+k-i) + \sum_{i=1}^{n} d_i \; x(n+k-i)$$

$$- \sum_{j=1}^{k-1} \alpha_j \sum_{i=1}^{j-1} b_i \; x(k+(j-1)-i) \qquad (11)$$

In vector matrix form we have

$$y(n+k) = [y(k)\cdots y(k+n-1)] \begin{bmatrix} -\alpha_n \\ \vdots \\ -\alpha_1 \end{bmatrix}$$

$$+ [x(k)\cdots x(k+n-1)] \begin{bmatrix} \beta_n \\ \\ \beta_1 \end{bmatrix} \qquad (12)$$

where the vector $\beta = [\beta_n, \cdots \beta_1]'$ may be seen to be

$$\begin{bmatrix} \beta_n \\ \cdot \\ \cdot \\ \cdot \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & & \\ -\alpha_1 & 1 & 0 & \cdots & \\ -\alpha_2 & \alpha_1 & 1 & 0 & \cdots \\ \vdots & & & & \\ -\alpha_{n-1} & \cdot & \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \qquad (13)$$

Consequently we have achieved our objective of writing (1) in the form (2).

More simply we can write (2) in the form

$$y(n) = \sum_{i=1}^{k} -\alpha_i \, y(n-i) + \eta(n) \qquad (14)$$

where the sequenc $\mu(n)$ is a correlated sequence.

Alternatively (14) and (1) can be put into the form

$$\begin{aligned} s(n+1) &= \Phi \, s(n) + b \, \eta(n) \\ y(n) &= c' \, s(n) \end{aligned} \qquad (15)$$

where Φ and c' are as defined in (6) and

$$b = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \qquad (16)$$

2.5 On the Transformation Between the Autoregressive and the Continuous Time Models.

Our concern here is with the transformation between the coefficients $\{\alpha_i\}$ in the canonical form, discrete time representation

$$\begin{aligned} s(n+1) &= \Phi\, s(n) + b\, \eta(n) \\ y(n) &= c'\, s(n) \end{aligned} \tag{1}$$

where the matrices Φ, b and c are as defined in Section 2.4 and the equivalent continuous time representation

$$\begin{aligned} \dot{q}(t) &= A\, q(t) + b\, x(t) \\ y(t) &= c'\, q(t) \end{aligned} \tag{2}$$

where the canonical form matrix A, and the vector b and c are as defined in Section 2.

The representation in (1) is derived from the representation

$$\begin{aligned} q(n+1) &= F\, q(n) + f\, x(n) \\ y(n) &= c'\, q(n) \end{aligned} \tag{3}$$

where

$$\begin{aligned} F &= \exp TA \\ \Phi &= BFB^{-1} \end{aligned} \tag{4}$$

Lemma 1* There exists a nonsingular transformation, T, such that

$$\Lambda = T\Phi T^{-1} \tag{5}$$

where

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \lambda_k \end{bmatrix} \tag{6}$$

That is, Λ is a diagonal matrix consisting of the roots of the characteristic polynomial of Φ. (All of the roots $\lambda_1 \cdots \lambda_k$ are assumed to be distinct.)

Lemma 2. The characteristic polynomial for Φ is

$$g(\lambda) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_n \tag{7}$$

Consequently once the $\{\alpha_i\}$ are known the elements λ_i of the matrix can be determined from (7).

Lemma 3. The nonsingular diagonal matrix Λ is similar to the matrix C in the sense

$$\Lambda = e^C \tag{8}$$

---

*Well-known mathematical results will be identified as lemmas and quoted without proof. A sufficient reference for the results employed in this section is Chapter (3), Coddington and Levenson (8).

where

$$C = \begin{bmatrix} \log \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \log \lambda_k \end{bmatrix} = \begin{bmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_k \end{bmatrix} \tag{9}$$

Lemma 4. For every matrix C and every matrix P

$$Pe^C P^{-1} = e^{PCP^{-1}} \tag{10}$$

We equate

$$PCP^{-1} = TA \tag{11}$$

which is motivated by (4), and use the identification in (9), the definition of (A) and lemmas (1) and (2) to get the characteristic polynomial

$$h(\lambda) = \mu^n + Ta_1 \mu^{n-1} + \cdots Ta_n = \prod_{i=1}^{k} (\mu - \mu_i) \tag{12}$$

for the matrix TA. Since the roots μ_i $i=1,2\cdots k$ are known from (9), the parameters $\{a_i\}$ are determined directly from (12).

To summarize: The parameters $a_1, \cdots a_k$ are obtained in the following manner.

(1) Estimate $\alpha_1, \cdots \alpha_k$ using the autoregressive scheme.

(2) Form the characteristic polynomial

$$f_\alpha(\lambda) = \lambda^k + \alpha_1 \lambda^{k-1} + \cdots \alpha_k \tag{13}$$

(3) Determine the roots $\lambda_1, \cdots \lambda_k$ of $f_\alpha(\lambda)$ (the eigenvalues of the time discrete system (1)).

(4) Then the characteristic polynomial for the time discrete system (2) revealing the continuous system parameters $a_1, \cdots a_k$ is given by

$$f_a(\mu) = \mu^k + a_1\mu^{k-1} + \cdots a_k = \prod_{i=1}^{k} (\mu - \log \lambda_i) \qquad (14)$$

2.6 Estimation in the Autoregressive Model

In the preceding section it was demonstrated that hte identification of the unknown parameters of a k/2 d.o.f. linear dynamical system excited by white noise could be associated with the estimation of the unknown parameters of an autoregressive scheme of order k. In this section we review the theory associated with the estimation of the parameters of the white noise residual autoregressive scheme

$$y(t) = \sum_{i=1}^{k} \alpha_i \, y(t-i) + e(t); \quad t = k+1, \cdots m \tag{1}$$

where $e(t)$, $t = 0 \pm 1 \cdots$ is a zero mean independent, identically distributed gaussian sequence with variance σ^2. The system is assumed to be observed over the finite duration interval specified by $t = k+1, \ldots m$.

For $t = k+1, \ldots m$ (1) can be written as

$$\begin{bmatrix} y(k+1) \\ \cdot \\ \cdot \\ \cdot \\ y(m) \end{bmatrix} = \begin{bmatrix} y(1) & \cdots & y(k) \\ & & \\ & & \\ \cdot & \cdots & y(m-1) \end{bmatrix} \begin{bmatrix} \alpha_k \\ \cdot \\ \cdot \\ \alpha_1 \end{bmatrix} + \begin{bmatrix} e(k) \\ \cdot \\ \cdot \\ \cdot \\ e(m) \end{bmatrix} \tag{2}$$

which is recognized to be in the least square parameter estimation form (see Quarterly Progress Report #2). In matrix form (2) is

$$Y_m = S_m{}'\alpha + e_m \tag{3}$$

where

$$y_m = \begin{bmatrix} y(k+1) \\ \vdots \\ y(m) \end{bmatrix} ; \quad S_m' = \begin{bmatrix} y(1) & \dots & y(k) \\ \vdots & & \\ & & y(m-1) \end{bmatrix} ;$$

$$\alpha = \begin{bmatrix} \alpha_k \\ \vdots \\ \alpha_1 \end{bmatrix} ; \quad e_m = \begin{bmatrix} e(k) \\ \vdots \\ e(m) \end{bmatrix} \tag{4}$$

The normal equations for the estimate, $\hat{\alpha}_m$, of the unknown parameter vector α (after m observations of $y(t)$) is

$$S_m y_m = S_m S_m' \hat{\alpha}_m \tag{5}$$

It is instructive to examine (5) in component form. This is given by

$$\begin{bmatrix} y(1) & y(2) & \cdots & y(m-k) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ y(k) & & \dots & y(m-1) \end{bmatrix} \begin{bmatrix} y(k+1) \\ \cdot \\ \cdot \\ \cdot \\ y(m) \end{bmatrix} =$$

$$\begin{bmatrix} y(1) & y(2) & \cdots & y(m-k) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ y(k) & & \dots & y(m-1) \end{bmatrix} \begin{bmatrix} y(1) & \cdots & y(k) \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ y(m-k) & & \cdots y(m-1) \end{bmatrix} \hat{\alpha}_m \tag{6}$$

Or

$$\begin{bmatrix} \hat{R}_0 \\ \\ \\ \\ \hat{R}_{k-1} \end{bmatrix} = \begin{bmatrix} \hat{R}_{0,0} & \hat{R}_{0,1} & \cdots & \hat{R}_{0,k-1} \\ \hat{R}_{1,0} & \hat{R}_{1,1} & \cdots & \vdots \\ \vdots & & & \vdots \\ \hat{R}_{k-1,0} & \cdots & \cdots & \hat{R}_{k-1,k-1} \end{bmatrix} \hat{\alpha}_m$$

where

$$\hat{R}_{r,s} = \frac{1}{m-k-1} \sum_{i=1}^{m-k-1} y(i+r)\, y(i+s) \tag{7}$$

and $\hat{R}_{r,s}$ is the estimator of the covariance $R(r-s) = E\{Y(t-r)Y(t-s)\}$ of the stationary process $\{Y(t),\ t = 0, \pm 1, \cdots\}$.

Following Hannan (4) and Anderson (9), the estimators can be seen to be asymptotically unbiased and normally distributed with covariance matrix

$$\mathrm{Cov}\ \hat{\alpha},\hat{\alpha} = (\mathrm{Cov}\ \hat{\alpha}_i,\hat{\alpha}_j) = \hat{\sigma}_k^{\ 2} \begin{bmatrix} \hat{R}_{0,0} & \cdots & \hat{R}_{0,k-1} \\ \vdots & & \vdots \\ \hat{R}_{k-1,0} & \cdots & \hat{R}_{k-1,k-1} \end{bmatrix}^{-1} \tag{8}$$

where $\hat{\sigma}_k^{\ 2}$ is the estimated residual variance given by

$$\hat{\sigma}_k^{\ 2} = \hat{R}(0) - \hat{\alpha}_1\hat{R}(1) - \cdots - \hat{\alpha}_k\hat{R}(k) \tag{9}$$

and where

$$\hat{R}(p) = \hat{R}_{j,j+p} \quad \text{for any } j. \tag{10}$$

It is interesting and useful to observe that by multiplying (1) by $y(t)$ and taking expectations we obtain

$$R(0) = \alpha_1 R(1) + \alpha_2 R(2) + \cdots \alpha_k R(k) + R_{ey}(0) \qquad (11)$$

Since the random variable $e(t)$ is assumed to be independent of the random variables $e(t-1), e(t-2), \cdots$ it is certainly independent of $y(t-1, y(t-2), \cdots$. Therefore in (11) we can substitute

$$R_{ey}(0) = R_{ee}(0) = \sigma^2_{ee} \qquad (12)$$

where σ^2_{ee} is the average "power" of the input process, $\{e(t), t=0, \pm 1, \cdots\}$. Since it is known (4) that the sample covariance matrix in (8) converges in probability to the true covariance matrix, for m large the residual variance (9) is an explanation of the extent to which the hypothesized model accounts for the observed power, $\hat{R}(0)$, where $R(0) = Ey(t)\, y(t)$. Consequently as k increases, the residual variance approaches the constant σ^2_{ee}, the input power.

Returning our attention to equations (1) - (8), for a given data set and m sufficiently large, one could determine a confidence region for the estimates $\hat{\alpha}_i$. Rather than pursue this point extensively, we note that the diagonal terms of the matrix in (8) designate the variance of the estimate of the corresponding estimate α_i. When the square root of each of these terms is significantly smaller than the estimate of α_i, the true value of α_i will well be in a region corresponding to any reasonable confidence coefficient.

In conjunction with the large sample procedure suggested to estimate the autoregressive coefficients, we wish to explore procedures to determine, k, the order of the hypothesized autoregressive model. Several alternatives are available for this purpose. One heuristic approach is to compute the residual variance statistic (9) for successive values of k.

That is, a simple practical hypothesis test is to compute the correlation matrix and the estimates α_i for as high an order k of regressive scheme that we are willing to consider. The estimate of k is sufficient if for no greater value of k is the residual variance, $\hat{\sigma}_k{}^2$, significantly decreased. Observe that this heuristic approach is an implicit application of the inspection scheme earlier. That is, the estimate of k suggested is the largest number k for which

$$|\hat{\alpha}_k| \gg \text{Cov } \hat{\alpha}_k, \hat{\alpha}_k. \tag{13}$$

There are a number of more formal alternatives available to test the order of the autoregressive model. Closely related is the test statistic

$$\lambda = \frac{\hat{\sigma}^2_{p+q}}{\hat{\sigma}^2_p} \tag{14}$$

which is used to distinguish between the hypotheses

$$H_0 \; ; \quad k = p+q \tag{15}$$

and

$$H_1 \; ; \quad k = p.$$

Whittle (10) and Anderson (9) have demonstrated that the related quantity ψ^2* is distributed, χ_q^2, i.e., chi-squared with q degrees of freedom, under the hypothesis H_0, where

$$\psi^2 = \left(\frac{1-\lambda}{\lambda}\right) (m-(p+q)) \sim \chi_q^2 \tag{16}$$

Similarly, hypothesis tests based on a statistic other than (14) (using partial and multiple correlation coefficients as well as spectral estimates) have been analyzed to determine the order of an autoregressive scheme. At this point the references by Hannan, Whittle, and Anderson (4), (5), (6),(10) provide sufficient reference and a bibliography for this topic.

---

*The test statistic ψ^2 is in the same form as the energy fit criterion suggestion in (2). It is in the form $\psi^2 = \hat{\sigma}_k{}^2/\hat{R}(0)$, where $\hat{\sigma}_k{}^2$ and $\hat{R}(0)$ are defined in (7), (9) and (10).

2.7 An Example

As a test of our computation programs and as an illustration of some of the material in this section we have considered the Kendall (11) autoregressive scheme

$$y(t) = 1.1\, y(t-1) - 0.5\, y(t-2) + \eta(t) \qquad (1)$$

A series of 100 gaussian independent unit variance samples was generated to correspond to the quantity $\eta(t)$ in (1). The recursive relationship in (1) was used to generate the sequence $\{y(t)\}$. From the $\{y(t)\}$ we compute the appropriate correlation function estimates and the corresponding normal equations appear in the form

$$\begin{bmatrix} \sum_{i=1}^{m-2} y(i)y(2+i) \\ \sum_{i=1}^{m-2} y(i+1)y(2+i) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m-2} y(i)y(i) & \sum_{i=1}^{m-2} y(i)y(1+i) \\ \sum_{i=1}^{m-2} y(i)y(1+i) & \sum_{i=1}^{m-2} y(i+1)y(i+1) \end{bmatrix} \begin{bmatrix} -\hat{\alpha}_2 \\ -\hat{\alpha}_1 \end{bmatrix} \qquad (2)$$

Equivalently we can write

$$\begin{bmatrix} \hat{R}_{0,2} \\ \hat{R}_{1,2} \end{bmatrix} = \begin{bmatrix} \hat{R}_{0,0} & \hat{R}_{0,1} \\ \hat{R}_{1,0} & \hat{R}_{1,1} \end{bmatrix} \begin{bmatrix} -\hat{\alpha}_2 \\ -\alpha_1 \end{bmatrix} \qquad (3)$$

Corresponding to the tabulated computer results in the pages immediately following, the solution of (2) gives the results

$$\begin{array}{cc} \alpha \text{ True} & \alpha \text{ Estimated} \\ 1.100 & 0.9898 \\ -0.500 & -0.4720 \end{array} \tag{4}$$

with the normalized covariance (correlation coefficient) matrix,

$$\hat{\rho} = \frac{\hat{R}_{i,i}}{\hat{R}_{0,0}} = \begin{bmatrix} 1.000 & 0.675 \\ 0.675 & 1.005 \end{bmatrix} \tag{5}$$

Additional computations for this example for larger values of m and for k=1 as well as for higher order systems are in progress.

```
M=100 K=  2 N=  2

INDEPENDENT GUASSIAN NOISE USED

ALPHA VECTORS OUTPUTTED IN FOLLOWING FORM

          ALPHA( I )
          ALPHA(I-1)
            . . .
          ALPHA( 1 )


     ALPHA
     0.5000
    -1.1000

        ALPHA HAT     DELTA1       DELTA2      PDELTA1      PDELTA2
          0.4720   0.4030E 02 -0.1226E 00 -0.4712E 00 -0.1213E-02
         -0.9898   0.1391E 03  0.2112E-01  0.9883E 00  0.9165E-03

MATRIX S * S TRANSPOSE, BY ROWS
ROW  1
 0.20612137E 03  0.13904459E 03
ROW  2
 0.13904459E 03  0.20704931E 03

MATRIX ( 1 / SST(1,1) ) * SST, BY ROWS
ROW  1
 0.10000000E 01  0.67457631E 00
ROW  2
 0.67457631E 00  0.10045020E 01

MATRIX ( S * S TRANSPOSE ) INVERSE, BY ROWS
ROW  1
 0.88695295E-02 -0.59563592E-02
ROW  2
-0.59563592E-02  0.88297781E-02

MATRIX ( ( S * S TRANSPOSE ) INVERSE ) * ( S * S TRANSPOSE ), BY ROWS
ROW  1
 0.99999999E 00 -0.22204461E-15
ROW  2
 0.11102230E-15  0.10000000E 01

THE U'S IN ORDER
 0.73880436E 00  0.25929414E 00  0.13433518E 00  0.35998884E 00 -0.10392822E 01  0.40533272E 00  0.11927098E 00  0.30288287E 00
-0.13604308E 01 -0.70509187E 00  0.27915216E 01  0.16523778E 01 -0.36634576E 00 -0.77068228E 00 -0.91616632E 00 -0.14337410E 01
 0.81943504E 00  0.89472724E 00 -0.10625780E 01  0.92535339E 00  0.90142372E 00 -0.94551990E 00  0.53437158E 00  0.41125434E 00
-0.43089435E 00  0.11876051E 01 -0.80677827E 00  0.49193846E 00  0.24058399E 01  0.39554541E-01  0.78293507E 00  0.43328656E 00
-0.80529159E 00 -0.64078411E 00  0.39908040E 00 -0.77706742E 00  0.42842595E 00  0.75012796E 00 -0.32582448E 00 -0.70510420E 00
-0.35498123E 00 -0.14292181E 01  0.34443318E 00  0.62212604E 00 -0.83025447E 00 -0.83467413E 00 -0.59377600E 00 -0.23945206E 00
-0.16325593E 01 -0.29061132E 00 -0.12369092E 00  0.31886029E 00 -0.91261391E 00 -0.11666644E 01  0.69248692E 00  0.22145173E-01
 0.51571644E 00  0.12215913E 00  0.43971628E-01 -0.80931842E 00  0.50691935E 00 -0.15767975E 01 -0.50110433E 00  0.14943577E 01
 0.79102659E-01  0.99825971E-02 -0.11568839E 01  0.45118502E 00  0.16265719E 01  0.17008390E 01  0.25469128E-01  0.16217687E 01
-0.28210331E 01  0.79451136E 00 -0.66368729E 00 -0.17378975E 01  0.15661460E 01 -0.14854908E 01  0.34543905E 00 -0.49665579E 00
-0.18320513E 00 -0.48399988E 00  0.77846043E 00 -0.16397269E 01  0.12427748E 01 -0.44114773E 00 -0.99774357E 00 -0.35647237E 00
```

```
-0.99989239E-02 -0.73097614E 00  0.10907307E 00 -0.66660363E 00 -0.66672965E 00  0.58952851E 00  0.57249165E 00 -0.29451326E 00
-0.56889378E 00  0.23259337E 00 -0.10099377E 01  0.10860424E 01

THE Y'S IN ORDER
-0.82408146E 00  0.80671754E-01  0.63511484E 00  0.10182793E 01 -0.23643243E 00 -0.36421259E 00 -0.16299665E 00  0.30569285E 00
-0.94267035E 00 -0.18948756E 01  0.11784936E 01  0.38961585E 01  0.33301818E 01  0.94443844E 00 -0.15423749E 01 -0.36025726E 01
-0.23722074E 01  0.86585438E-01  0.21876968E 00  0.11227073E 01  0.20270169E 01  0.72284504E 00  0.31599266E 00  0.39742374E 00
-0.15172455E 00  0.82199620E 00  0.17327984E 00  0.27154818E 00  0.26179029E 01  0.27834737E 01  0.25358046E 01  0.18309348E 01
-0.59165649E-01 -0.16213337E 01 -0.13548038E 01 -0.14556848E 01 -0.49652538E 00  0.93229242E 00  0.94795987E 00 -0.12849455E 00
-0.97030518E 00 -0.24323065E 01 -0.18459514E 01 -0.10226720E 00 -0.11877270E 00 -0.86919050E 00 -0.14904992E 01 -0.14444059E 01
-0.24761562E 01 -0.22921802E 01 -0.14070110E 01 -0.82761714E-01 -0.30014629E 00 -0.14554444E 01 -0.75842880E 00 -0.84404293E-01
 0.80208612E 00  0.10466560E 01  0.79425017E 00 -0.45897123E 00 -0.39507408E 00 -0.17818934E 01 -0.22636500E 01 -0.10471062E 00
 0.10957460E 01  0.12676585E 01 -0.31033257E 00 -0.52401005E 00  0.12053272E 01  0.32887039E 01  0.30403798E 01  0.33218345E 01
-0.68720506E 00 -0.16223315E 01 -0.21046493E 01 -0.32418461E 01 -0.94755997E 00 -0.90688369E 00 -0.17835301E 00 -0.23940226E 00
-0.35737111E 00 -0.75740696E 00  0.12399832E 00 -0.11246252E 01 -0.56312177E-01  0.59221491E-01 -0.90444384E 00 -0.13809713E 01
-0.10768455E 01 -0.12250205E 01 -0.70002671E 00 -0.82412276E 00 -0.12227513E 01 -0.34343656E 00  0.80608708E 00  0.76390082E 00
-0.13164642E 00 -0.29416810E 00 -0.12676993E 01 -0.16044282E 00

THE STARTING U'S IN ORDER
-0.76843097E-01 -0.13712071E 01

THE STARTING Y'S IN ORDER
-0.76843097E-01 -0.14557344E 01
```

3. REFERENCES

1. Quarterly Progress Report, No. 1. MASC for Goddard Space Flight Center, Greenbelt, Maryland, Contract No. NAS5-10106.

2. Quarterly Progress Report, No. 2.

3. Mann, H. B. and Wald, A., "On the Statistical Treatment of Linear Stochastic Equations", Econometrica, Vol. 11, 173-220. (1943)

4. Hannan, E. J., "Time Series Analysis", Methuen and Co., London (1960).

5. Whittle, P., "Prediction and Regulation", Princeton; Van Nostrand (1963).

6. Blackman, R. B., "Data Smoothing and Prediction", Addison-Wesley Publishing Co., Inc., (1965).

7. Lee, R. C. K., "Optimal Estimation, Identification and Control", Research Monograph No. 24, M.I.T. Press, Cambridge, Mass. (1964).

8. Coddington, E. A. and Levenson, N., "Theory of Ordinary Differential Equations", McGraw-Hill Book Co., (1955).

9. Anderson, T. W., "Determination of the Order of Dependence in Normally Distributed Time Series", Time Series Analysis, edited by M. Rosenblatt, Wiley (1963).

10. Whittle, P., "Tests of Fit in Time Series", Biometrica, Vol. 39, (1952).

11. Kendall, G. M., "On the Analysis of Oscillatory Time Series", J. Roy. Stat. Soc. A., Vol. 108, (1945).